

Integration of QSAR and *in Vitro* Toxicology

Martin D. Barratt

Unilever Environmental Safety Laboratory, Sharnbrook, Bedford,
United Kingdom

The principles of quantitative structure–activity relationships (QSAR) are based on the premise that the properties of a chemical are implicit in its molecular structure. Therefore, if a mechanistic hypothesis can be proposed linking a group of related chemicals with a particular toxic end point, the hypothesis can be used to define relevant parameters to establish a QSAR. Ways in which QSAR and *in vitro* toxicology can complement each other in development of alternatives to live animal experiments are described and illustrated by examples from acute toxicological end points. Integration of QSAR and *in vitro* methods is examined in the context of assessing mechanistic competence and improving the design of *in vitro* assays and the development of prediction models. The nature of biological variability is explored together with its implications for the selection of sets of chemicals for test development, optimization, and validation. Methods are described to support the use of data from *in vivo* tests that do not meet today's stringent requirements of acceptability. Integration of QSAR and *in vitro* methods into strategic approaches for the replacement, reduction, and refinement of the use of animals is described with examples. — *Environ Health Perspect* 106(Suppl 2):459–465 (1998). <http://ehpnet1.niehs.nih.gov/docs/1998/Suppl-2/459-465barratt/abstract.html>

Key words: QSAR, *in vitro*, alternative methods, skin corrosivity, eye irritation, validation, prediction model

Introduction

The principles of quantitative structure–activity relationships (QSAR) are based on the premise that the properties of a chemical are implicit in its molecular structure. Therefore, if a mechanistic hypothesis can be proposed linking a group of related chemicals with a particular toxic end point, the hypothesis is used to define relevant parameters to establish a structure–activity relationship. The resulting model is then tested and the hypothesis and parameters refined until an adequate model is obtained. These principles have been successfully applied in this laboratory to predict a skin permeability coefficients (1), the skin corrosivity of organic acids, bases, phenols (2,3) and electrophilic organic chemicals (4), and the eye irritation potential of neutral organic chemicals (5,6).

For a QSAR to be valid and reliable, the dependent property for all the chemicals covered by the relationship must be elicited by a mechanism that is both common and relevant to that dependent property. Attempts to derive QSARs for data sets in which either the dependent property is derived by more than one mechanism or the mechanism of action is wrongly defined do not usually lead to robust models.

The same principles that are applied to the development of QSARs must also be applied to the development of *in vitro* alternatives to animal tests if those methods are to be reliable. These principles have been overlooked in many cases, particularly in the prediction of acute toxic effects, with inevitable results. Some alternative tests determine end points that are

substantially different from those that they claim to predict because the mechanism modeled by the *in vitro* alternative represents only part of that which is active *in vivo*. In other cases, tests have been developed that can predict end points accurately for some classes of chemicals but are then wrongly assumed to be applicable to all chemical classes. The fact that different types of chemicals may elicit changes in a particular biological end point through different mechanisms clearly has not been appreciated.

This paper describes some of the ways in which QSAR and *in vitro* toxicology can complement each other in the development of alternatives to live animal experiments by using examples from acute toxicological end points. The mechanistic approach to QSAR and *in vitro* methods is examined in the context of assessing the mechanistic competence, improving the design of *in vitro* assays, and developing prediction models. The nature of biological variability is explored together with implications for the selection of sets of chemicals for test development, optimization, and validation. Methods are described in which QSAR can be applied to support the use of data from *in vivo* tests that do not meet today's stringent requirements of acceptability. Examples are given of the integration of QSAR into strategic approaches for the replacement, reduction, and refinement of the use of animals.

The Mechanistic Approach

Cases in which the mechanism of action is known or can be postulated provide sound arguments in favor of a mechanistic approach to the problem of designing alternative methods to replace animal experiments. Based on the premise that the properties of a chemical are implicit in its chemical structure, if the mechanistic basis for a specific toxicological property of a group of related chemicals can be elucidated and the relevant parameters measured or calculated, then, in principle, an *in vitro* alternative can be established. For an *in vitro* alternative to be valid and reliable, the specific toxicological property for all the chemicals covered by the alternative method must be elicited by a mechanism that is both common and relevant to that property. Attempts to derive *in vitro* alternatives for sets of chemicals in which either the toxicological property is derived by more

This paper was prepared as background for the 13th Meeting of the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC): Alternative Testing Methodologies held 26–31 January 1997 in Ispra, Italy. Manuscript received at EHP 9 May 1997; accepted 23 July 1997.

The author thanks numerous colleagues at the Unilever Environmental Safety Laboratory and at LHASA UK for their collaboration, particularly M. Chamberlain for his encouragement and continued intellectual support in the development of the ideas in this paper.

Address correspondence to Dr. M.D. Barratt, Marlin Consultancy, 10 Beeby Way, Carlton, Bedford MK43 7LW, UK. Telephone: 44 1234 720756. Fax: 44 1234 721533. E-mail: martin.d.barratt@btinternet.com

Abbreviations used: C3, C6, C8, C12, alkyl chain lengths of 3, 6, 8, 12 carbons; DEREK, deduction of risk from existing knowledge (a knowledge-based expert system); EC, European Community; ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals; ECVAM, European Centre for Validation of Alternative Methods; GLP, Good Laboratory Practice; IVSCT, *in vitro* skin corrosivity test; LLNA, local lymph node assay; logP, log(octanol/water partition coefficient); pK, log(ionization constant); QSAR, quantitative structure–activity relationship; Ry, Rz, minor inertial axes of a chemical.

than one mechanism or the mechanism of action is wrongly defined probably will not lead to a successful outcome.

A mechanistic understanding of the biological process underlying an *in vivo* toxicity assay to be replaced, therefore, can be expected to lead to identification of the parameters most appropriate to model the toxicity. The underlying scientific basis of alternative methods based on these principles should improve their credibility and assist their acceptance by regulatory authorities. It could be argued that the fact that regulatory acceptance often is not readily forthcoming is based partly on relatively poor understanding of the scientific (mechanistic) basis of many alternative methods.

In constructing a QSAR model the parameters, where possible, should be selected on the basis of an understanding of the mechanism of the process. When a mechanism is not understood sufficiently to be used to define appropriate parameters, an alternative procedure is to compute a large number of parameters and attempt to establish a statistical relationship with a few of those parameters. Parameters from such a QSAR may be useful in understanding the mechanistic basis of the process. Parameters used in the design of an *in vitro* model should also be selected on the basis of mechanistic understanding; selection of appropriate parameters can be aided by constructing a QSAR model using the same training data.

The chemicals used to construct a QSAR or *in vitro* method (the training set) should be selected on the basis of a common mechanism of action if possible and should adequately cover the parameter space in terms of dependent and independent variables. It is also important to realize that the predictive domain of both types of models is restricted to the same parameter space covered by the training set, i.e., the model can be used for interpolation but not for extrapolation.

One limitation that may apply to QSAR models more than to *in vitro* alternatives is that to date the successful application of QSAR has been restricted to modeling properties of pure chemicals. Concerning *in vitro* alternatives, it may be possible to construct a complete, or mechanistically competent, *in vitro* model and calibrate and validate it using mixtures of chemicals. If the constraints described above relating to mechanistic validity and parameter space are taken into account, there appears to be no reason why such models should not provide useful predictions.

Assessing Mechanistic Competence of *in Vitro* Tests

It has long been recognized that for a chemical to be biologically active it first must be transported from its site of administration to its site of action and then bind to or react with its receptor or target (7), i.e., biological activity is a function of partition and reactivity. If any QSAR or *in vitro* model is deficient in modeling either partition or reactivity, only a partial correlation with the *in vivo* response is likely to be observed. An example is the varying degrees of partial correlation with *in vivo* data found with the many *in vitro* methods developed and advocated as alternatives to the Draize rabbit eye irritation test (8). Thus, it follows that for an *in vitro* test to reliably predict *in vivo* toxic potential, it should be sensitive to the same parameters responsible for the effects *in vivo*; such a test would be expected to show a high degree of correlation with the response *in vivo*.

In vitro tests can be categorized as

- Empirical: those for which no mechanistic basis linking the *in vivo* and *in vitro* end points has been identified, for example, the pollen tube growth inhibition test for eye irritants (9).
- Mechanistic: those for which the mechanistic link is clearly identified, for example, *in vitro* photobinding to human serum albumin test for photoallergens (10).
- Analogous: those in which all or part of the *in vivo* system is reproduced *in vitro*, for example, the isolated rabbit eye test (11) and the *in vitro* skin corrosivity test (12). (It is implicit that mechanisms relevant to *in vivo* toxicity operate in an analogous assay.)

A hypothetical example of a deficient alternative test has been simulated (13) by omitting a key parameter from a QSAR model for the eye irritation potential of 46 neutral organic chemicals (5). In the original model, the parameters modeling partition are log[octanol/water partition coefficient], logP, a measure of hydrophobicity/hydrophilicity, and the minor principal inertial axes R_y and R_z representing the cross-sectional area of the molecule; the reactivity parameter is modeled by the computed dipole moment of each chemical.

Figure 1A is a plot of the first two principal components of all four parameters for the 46 chemicals. Except for one chemical, the plot discriminates completely between chemicals classified as eye irritants and those classified as nonirritants. In Figure 1B, the principal components are

replotted with the dipole moment (reactivity parameter) omitted. Although there is still some discrimination between irritant and nonirritant chemicals—probably because there is a partial correlation between logP and dipole moment—there now is one area in the plot where irritants and nonirritants overlap considerably. The molecular parameters that remain in the model are logP, and the R_y and R_z modeling partitions; however, because dipole moment is absent, full assessment of biologic activity cannot be made.

Using mechanistically based QSAR techniques has profound implications for assessment of the practical utility of *in vitro* tests. Elucidating the putative mechanism of action for a class of chemicals facilitates evaluation of an *in vitro* method for predicting the *in vivo* toxic potential of an untested chemical. If the QSAR is mechanistically based, its independent variables should define the mechanistic requirements of the *in vitro* assay as well as the scope and limitations of its parameter space. Therefore, using these mechanistic requirements there is a greater likelihood of determining (or designing) an appropriate *in vitro* assay.

Chemical parameters relevant to the mechanism of action must be identified for QSAR development. Once developed, the QSAR can be used to test the mechanistic hypothesis developed earlier. Significant outliers may indicate that mechanisms outside the existing QSAR are operating, so new QSARs may need to be developed. Eventually, a robust QSAR model is developed. Similarly, the process for developing an *in vitro* assay should start with defining the toxicologic phenomenon to be modeled or replaced, defining relevant mechanism, and categorizing chemicals on the basis of mechanism of action. An end point in the *in vitro* assay should be selected on the basis of relevance to the *in vivo* toxicity.

A second pitfall of *in vitro* alternatives through which they may and a way in which fall short of *in vivo* systems they purport to replace is the problem of multiple mechanisms. As noted earlier, some alternative tests have been developed that can predict end points accurately for some classes of chemicals but are often wrongly assumed to be applicable to all chemical classes. Chemicals eliciting a biologic response by mechanisms other than those covered by the scope of the training set will appear as significant outliers. That different types of chemicals elicit changes in particular biologic end points through different mechanisms often is not recognized. If no mechanism is identified,

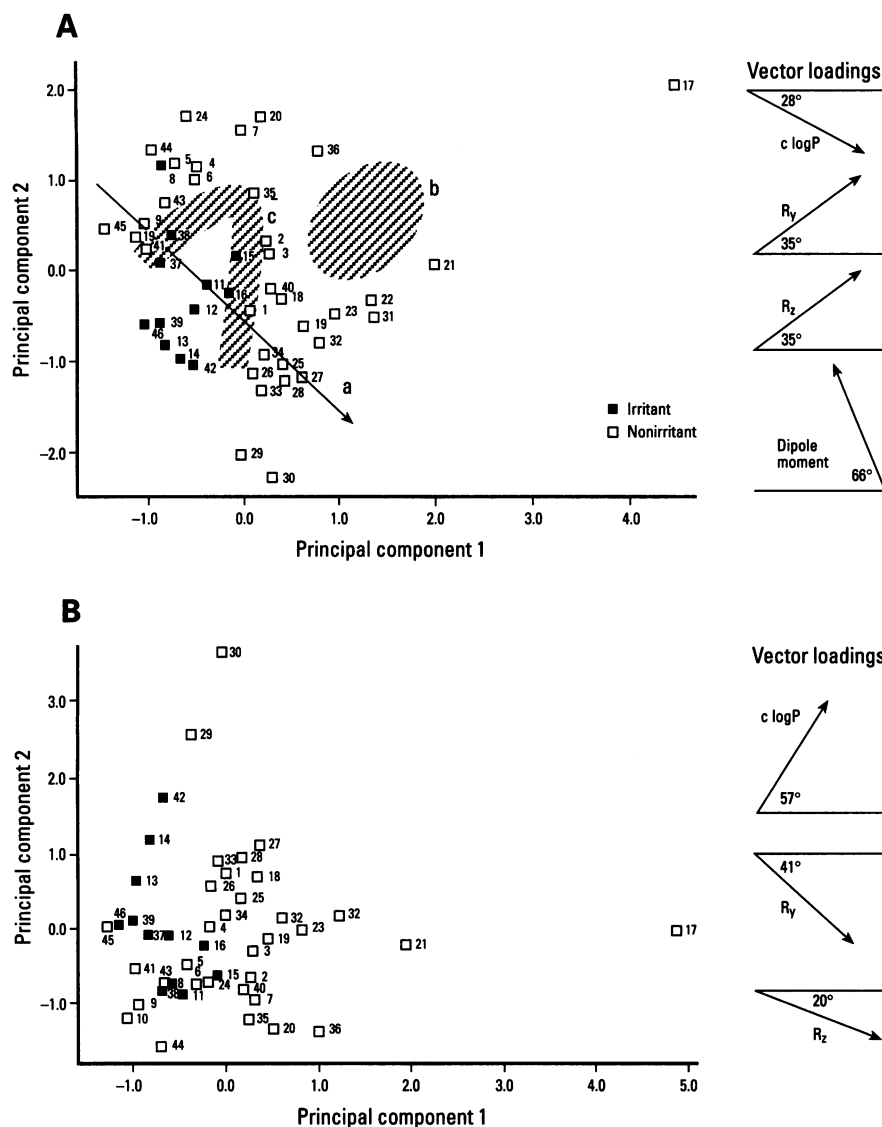


Figure 1. (A) Plot of the first two principal components of $c \log P$, R_y , and R_z , and dipole moment for 46 neutral organic chemicals showing vector loadings. Key: 1, 2-ethyl-hexan-1-ol; 2, methyl trimethyl acetate; 3, ethyl trimethyl acetate; 4, *n*-butyl acetate; 5, ethyl acetate; 6, cellosolve acetate; 7, ethyl-2-methylacetoacetate; 8, methyl acetate [methyl acetate is thought to be an eye irritant through a mechanism not covered by this QSAR (5)]; 9, propylene glycol; 10, glycerol; 11, isopropanol; 12, isobutanol; 13, *n*-butanol; 14, *n*-hexanol; 15, butyl cellosolve; 16, cyclohexanol; 17, 4,4-methylene-bis(2,6-di-*t*-butylphenol); 18, 4-bromophenol; 19, xylene; 20, 3-chloro-4-fluoronitrobenzene; 21, 1,3-diisopropylbenzene; 22, 1-methylpropylbenzene; 23, 3-ethyltoluene; 24, 2,4-difluoronitrobenzene; 25, styrene; 26, toluene; 27, 3-methylhexane; 28, 2-methylpentane; 29, 1,9-decadiene; 30, dodecane; 31, 1,5-dimethylcyclooctadiene; 32, *cis*-cyclooctene; 33, methylcyclopentane; 34, 1,5-hexadiene; 35, methyl isobutyl ketone; 36, methyl amyl ketone; 37, methyl ethyl ketone; 38, acetone; 39, allyl alcohol; 40, chloroform; 41, 2-methoxyethanol; 42, octan-1-ol; 43, dimethylformamide; 44, dimethylsulfoxide; 45, formamide; 46, 2(2-ethoxyethoxy)ethanol. (B) Plot of the first two principal components of $c \log P$, R_y , and R_z for 46 neutral organic chemicals showing vector loadings, i.e., dipole moment omitted. Key is identical to that for Figure 1A. Reproduced from Chamberlain and Barratt (13) with permission.

QSAR methods may help elucidate the mechanisms operating in *in vitro* assays.

Use of QSARs to Develop and Refine Prediction Models

Prediction models are algorithms that use results from an alternative method to predict toxicity observed *in vivo*. A prediction model

must define the relationship between all possible results from an *in vitro* test with the *in vivo* end point, both qualitatively and quantitatively, including confidence limits; it also must define the scope of the alternative test in terms of the physical and chemical properties of the substances for which it is valid, i.e., it must define the

physicochemical parameter space for which it is valid (14).

A QSAR can be regarded as a prediction model and predictions can be made as long as the chemical parameters responsible for the toxicity of the chemical in question are within the chemical parameter space of the model (1,13).

An example in which QSAR was used as a prediction model to construct a hypothesis that was subsequently tested is provided by Whittle et al. (15). This work was based on a QSAR for the corrosivity of organic acids (2), with the putative mechanism that corrosivity is a function of the ability of the chemical to permeate the skin together with its cytotoxicity expressed in this case as acidity (pK_a). The study examined the corrosive potential of a series of fatty acids ranging from propanoic acid (C3) to dodecanoic acid (C12) using the *in vitro* skin corrosivity test (IVSCT) (16). In this series of fatty acids, the cytotoxicity parameter, pK_a remains constant; changes in skin corrosivity potential are determined therefore entirely by the variables that model skin permeability— $\log[\text{octanol/water partition coefficient}]$ ($\log P$), molecular volume, and melting point (17). Because a number of chemicals when tested on human skin have shown significantly different corrosivity results than when tested on animal skin (12,18), this series of fatty acids was investigated in the IVSCT using both rat skin and human skin. A principal components map (ex ref. 16) illustrating the corrosivity of organic acids is shown in Figure 2. The corrosivity/irritation profile of the fatty acids series toward rat skin *in vitro* is identical to that toward rabbit skin *in vivo*.

All fatty acids with alkyl chain lengths up to and including C8 were found to be corrosive to rat skin. When human skin was used, the corrosive/noncorrosive threshold was shifted to around the C6 fatty acid. The mechanistic interpretation of these results is consistent with the greater permeability barrier known to be associated with human skin compared with that of rat skin. This particular example also illustrates that animals are not necessarily good models for humans. The power of the QSAR as a predictive model lies in the ability to identify when the rat may not be a good model for humans in terms of chemical parameter space. Although rat skin might be useful for predicting corrosivity toward human skin for some chemicals, there are specific chemicals for which the rat would overpredict the corrosive hazard for human skin.

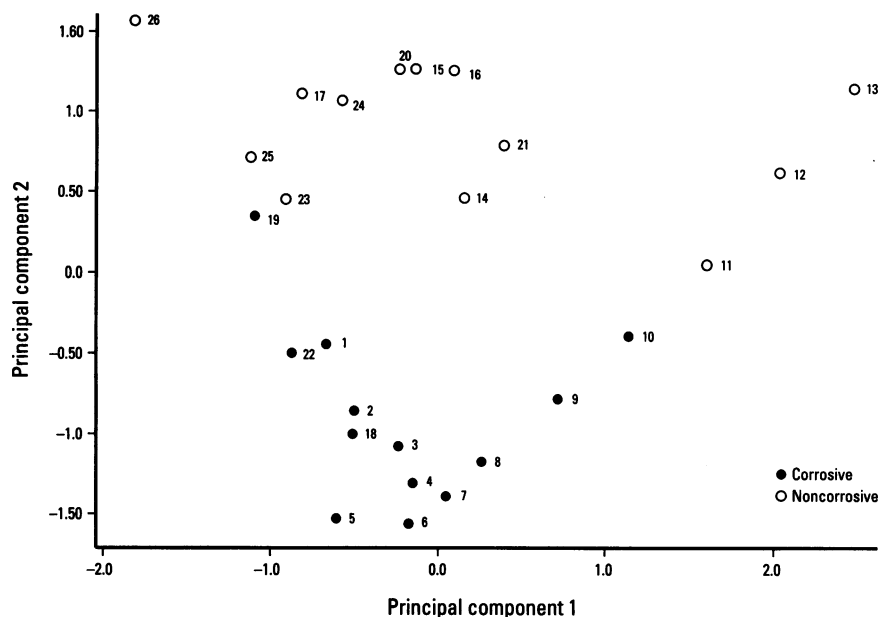


Figure 2. Plot of the first two principal components of logP, molecular volume, melting point, and pK_a for 27 organic acids. Key: 1, dichloroacetic acid; 2, bromoacetic acid; 3, mercaptoacetic acid; 4, acrylic acid; 5, formic acid; 6, acetic acid; 7, propanoic acid; 8, butanoic acid; 9, hexanoic acid; 10, octanoic acid; 11, decanoic acid; 12, dodecanoic acid; 13, tetradecanoic acid; 14, benzoic acid; 15, 2,4,6-trichlorophenol; 16, 2-bromobenzoic acid; 17, 4-nitrophenylacetic acid; 18, citric acid; 19, lactic acid; 20, oxalic acid; 21, salicylic acid; 22, *trans*-cinnamic acid; 23, cyanoacetic acid; 24, malonic acid; 25, succinic acid; 26, maleic acid; 27, sulfamic acid. Reproduced from Whittle et al. (16) with permission.

Predictions made in parameter space about which there is little knowledge can be tested by conducting new experiments in an optimum manner rather than by incremental expansion or repeated testing over small regions of parameter space.

Understanding Biologic Uncertainty

A recurring problem encountered in QSAR models when classifying toxicologic hazards is that of biologic uncertainty at boundary regions. The concept of the boundary region is based on the fact that most regulatory schemes operate initially by quantizing continuous biological (toxicological) data into discrete hazard bands that can be used conveniently in the regulatory process. It is the biologic variability inherent in toxicologic testing that leads to uncertainty in classifying boundary regions. One example of a boundary region is illustrated in Figure 1A by the area marked "c". The position of a chemical on a principal components plot is determined solely by its molecular features, whereas assignment of a toxicity classification depends entirely on the results from a variable biological assay, in this case the Draize rabbit eye test. Two well-conducted Draize rabbit eye irritation tests on the same chemical, for example, could lead to a nonirritant

classification for one and an irritant classification for the other. Away from the boundary region, the inherent biological variability is less likely to result in two tests leading to different classifications.

Other examples of such variability have been cited previously. For example, triethanolamine, furfurylamine, *N*-methylmorpholine, and *N*-ethylmorpholine are labeled as irritants by some suppliers and as corrosives by others; these chemicals lie in a region of a principal components plot for the skin corrosivity of organic bases in which the distinction between corrosive and noncorrosive is unclear (3). QSAR techniques such as principal components analysis allow one to visualize and hence to predict regions of chemical parameter space in which ambiguity in *in vivo* results may arise.

Another source of variability in the classification of chemicals may be in regulatory classification schemes. The European Community (EC) classification scheme for skin irritants (19), uses two different scoring systems, depending on whether the test has been carried out using three or more than three animals. For a three-animal test, the classification is based on two or more animals reaching the threshold score; when more than three animals are tested, the classification is based on the average score

calculated over all the animals tested. With this scoring system a chemical with skin irritation potential on the irritant/nonirritant threshold is less likely to be classified as a skin irritant if it is tested on more than three animals because it is possible for a single animal with a low irritancy score to reduce the average score to below the threshold even if the individual scores of all the other animals are above the threshold.

Selection of Sets of Test Chemicals for Validation of *in Vitro* Tests

The performance of an *in vitro* test is assessed by its capability of correctly predicting the *in vivo* response. It is also reasonable to propose that the *in vitro* test would not predict the result of an *in vivo* test any more accurately than a repeat *in vivo* test.

There are three ways to select a set of test chemicals to validate *in vitro* tests. First, a set of chemicals could be selected for which the biological data are either all very high or all very low responses. Such a set of chemicals probably would lead to a very optimistic, although misleading, outcome for the validation study but would do little to test the discriminating power of the *in vitro* test. Second, if the set of test chemicals consisted entirely of chemicals with uncertain *in vivo* data, confidence on the *in vitro* predictions would at least be equally uncertain. The third and most scientific approach is to select a set of chemicals for which the *in vivo* responses cover the whole range of biological responses—and QSAR affords a way of doing this (13).

This third method has been demonstrated using the principal components map for the eye irritation potential of neutral organic chemicals (5). Using the principal components map allows selection of chemicals that cover the widest possible parameter space in terms of both biological activity and physicochemical properties. For example, this may be achieved by selecting a series of chemicals that would start in an area predicted to be nonirritant, pass through the irritant area, and move out again into the nonirritant area. This is illustrated by track "a" in Figure 1A. The same principal components map can also be used to identify regions of parameter space incompletely covered by the current database (e.g., in the region marked "b" in Figure 1A). Obtaining biological test data from chemicals in these regions would be essential for the completeness of the nonanimal model. Similarly, the map can be used to identify regions of parameter

space well covered by the current database. Testing additional chemicals in these regions would add little valuable information and therefore might be a waste of valuable resources.

Similar techniques have been used in connection with the selection of test chemicals for the study sponsored on skin corrosivity sponsored by the European Centre for Validation of Alternative Methods (ECVAM) (20,21).

Availability of Good Quality Data

A major challenge facing researchers developing either *in vitro* models or QSARs is the sparse availability of high-quality data from experiments with animals (22). However, where there are biological data that do not meet today's stringent requirements of acceptability (see above), particularly historical data generated prior to the advent of Good Laboratory Practices (GLP), it is possible that QSARs may be used to validate these data for use in alternative tests. If QSAR techniques can be used to demonstrate that the results of these tests are consistent with the physicochemical attributes of the chemicals when compared with the results from tests conforming to current acceptance criteria, they can be deemed acceptable to use for development and validation of *in vitro* alternative methods.

In a recent QSAR study of the eye irritation potential of neutral organic chemicals (6), a relationship was established between the eye irritation data of the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) data bank (23) and EC classifications derived from those data and a large body of historical eye irritation data assessed by the criteria of Carpenter and Smyth (24); the utility of the latter data for use in nonanimal alternatives is therefore increased.

Integrated Approaches

One of the strengths of QSAR is the ease with which it can model partition either directly, e.g., logP, or through a combination of parameters, as exemplified by the modeling of skin permeability (1,25). On the other hand, one of the major strengths of *in vitro* toxicology is the ease of measuring cytotoxicity parameters, properties that depend more on the reactivity properties of the chemical and less on partition. Although the reactivity parameters of chemicals can be calculated using chemical modeling software or even measured directly, these methods are often comparable only within relatively small

areas of chemistry. A practical solution to this problem is to use cytotoxicity data from *in vitro* toxicology techniques as the independent variables for reactivity in QSARs. An example of this approach currently under development is the use of neutral red uptake data to measure the cytotoxicity of electrophilic organic acids. These data combined with parameters used in a previous QSAR study of the corrosivity of organic acids (2) are proving useful in discriminating between chemicals with the EC classifications R34 (corrosive, causes burns) and R35 (corrosive, causes severe burns) (26).

An even more important contribution for QSAR is to play a role in integrated strategies that lead to a reduction in the use of experimental animals. Two examples of proposed strategies are described briefly.

A scheme illustrating the strategic approach to skin sensitization hazard identification is illustrated in Figure 3 (27). In the first instance, a substance of defined chemical structure to be investigated is entered into the DEREK (deduction of risk from existing knowledge) expert system (28,29) to determine if it contains a structural alert—a fragment of chemical structure that could lead to the reactivity component of skin sensitization (30). If no structural alert is identified, the chemical is not likely to be a significant skin sensitizer; however, this should be confirmed using a standard animal assay—the mouse local lymph node assay (LLNA) (31) is considered most appropriate. If a skin sensitization structural alert is identified, the chemical has met the first of the two criteria for classification.

To be classified as a skin sensitizer, a chemical must be able not only to react chemically with skin protein either directly or after appropriate metabolism but also to partition into the relevant skin compartment. Skin permeability is assessed using a QSAR model (1). If the skin penetration of the chemical is sufficiently high, the chemical is assumed also to have significant skin sensitization potential and can be classified and labeled accordingly. If skin penetration is judged to be insignificant, the chemical is considered unlikely to be a skin sensitizer. In the latter case and where the extent of skin penetration is judged to be moderate or equivocal, it is considered advisable to assess the chemical with a suitable animal model. Again, the LLNA is considered the most appropriate test method.

Regardless what triggers the decision to conduct an LLNA, the practical outcome is the same. If the chemical is positive in this

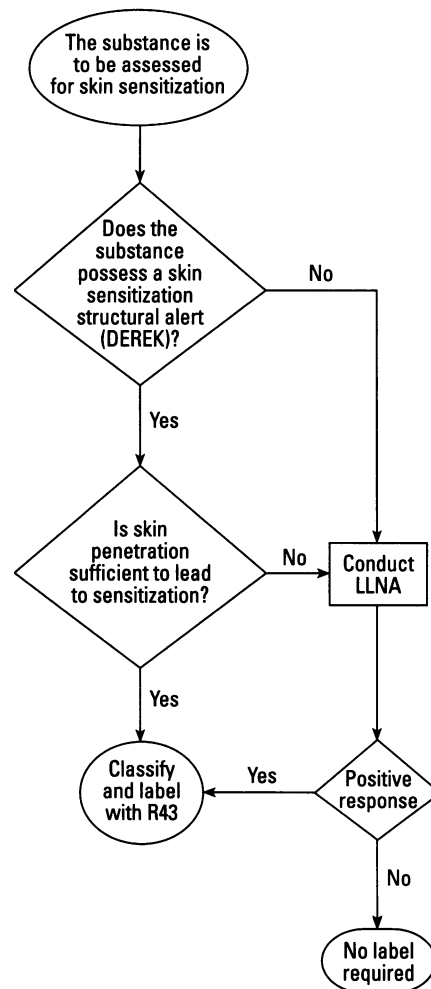


Figure 3. A strategic approach to skin sensitization hazard identification. Reproduced from Basketter et al. (27) with permission.

assay, it should be classified a skin sensitizer and labeled accordingly (32). When the result clearly does not meet the criteria for classification, then in our view no further work should be necessary. The chemical may be regarded as having insufficient sensitization potential to merit classification and labeling as a skin sensitizer. This proposed strategy provides an important opportunity for both substantial reduction and refinement of animal usage in a manner that does not compromise the existing standards of classification and labeling of skin sensitization hazard in the European Union.

The second strategy for identifying and classifying chemicals causing skin irritation or corrosivity uses a combination of QSAR and *in vitro* methods (33). In cases in which the chemicals are predicted confidently to be noncorrosive and ethical approval can be obtained, the final stage in the strategy is a

human 4-hr patch test. A practical example of the use of this strategy has already been provided in "Use of QSARS to Develop and Refine Prediction Models."

Conclusions

A number of ways to integrate QSAR knowledge and *in vitro* method development and evaluation have been explored. A major limitation is the mechanistic understanding of many toxicological phenomena. Where such understanding does not already exist, we have shown how development of QSARs may help elucidate and clarify such mechanisms of action. This approach requires careful and systematic

consideration of toxic effects, especially reexamining the descriptors of *in vivo* toxicity. With the current state of knowledge, the simpler toxicities, such as eye and skin irritation and skin sensitization, are more amenable to QSAR analysis. There already are indications, however, that more complex toxicities, for example $\alpha_2\mu$ -globulin nephropathy (34), peroxisome proliferation (35), and teratogenicity (36), may be similarly amenable to these approaches.

Ultimately, QSAR models have the potential to be used reliably to predict toxicity of chemicals and thus replace animals in some toxicity studies. In the meantime, QSAR methods can be used to

- evaluate the mechanistic competence of *in vitro* methods
- refine existing *in vitro* methods and give insight into the design of new ones
- reduce the need for testing and hence reduce the number of animals used
- provide balanced selections of chemicals for use in the validation of *in vitro* tests
- check the acceptability and consistency of biological data for use in the development and validation of alternative methods
- play a role in integrated strategies along with both *in vitro* and *in vivo* tests and eventually lead to a reduction in the use of experimental animals.

REFERENCES AND NOTES

1. Barratt MD. Quantitative structure activity relationships for skin permeability. *Toxicol in Vitro* 9:27-37 (1995).
2. Barratt MD. Quantitative structure activity relationships for skin corrosivity of organic acids, bases and phenols. *Toxicol Lett* 75:169-176 (1995).
3. Barratt MD. QSARs for skin corrosivity of organic acids, bases and phenols: principal components and neural network analysis of extended datasets. *Toxicol in Vitro* 10:85-94 (1996).
4. Barratt MD. Quantitative structure activity relationships for the skin irritation and corrosivity of neutral and electrophilic organic chemicals. *Toxicol in Vitro* 10:247-256 (1996).
5. Barratt MD. A quantitative structure activity relationship for the eye irritation potential of neutral organic chemicals. *Toxicol Lett* 80:69-74 (1995).
6. Barratt MD. QSARs for the eye irritation potential of neutral organic chemicals. *Toxicol in Vitro* 11:1-8 (1997).
7. Dearden JC. Physico-chemical descriptors. In: *Practical Applications of Quantitative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology* (Karcher W, Devillers J, eds). Dordrecht, The Netherlands: Kluwer, 1990;25-59.
8. Bagley DM, Bruner LH, de Silva O, O'Brien KAF, Uttley M, Walker AP. An evaluation of five potential alternatives *in vitro* to the rabbit eye irritation test *in vivo*. *Toxicol in Vitro* 6:275-284 (1992).
9. Kristen V, Hoppe V, Pape W. The pollen tube growth test: a new alternative to the Draize eye irritation test. *J Soc Cosmet Chemists* 44:153-162 (1993).
10. Barratt MD, Brown KR. Photochemical binding of photoallergens to human serum albumin; a simple *in vitro* method for screening potential photoallergens. *Toxicol Lett* 24:1-6 (1985).
11. Burton ABG, York M, Lawrence RS. The *in vitro* assessment of severe eye irritants. *Food Chem Toxicol* 19:471-480 (1981).
12. Oliver GJA, Pemberton MA. The identification of corrosive agents for human skin *in vitro*. *Food Chem Toxicol* 24:513-515 (1986).
13. Chamberlain M, Barratt MD. Practical applications of QSAR to *in vitro* toxicology illustrated by consideration of eye irritation. *Toxicol in Vitro* 9:543-547 (1995).
14. Bruner LH, Carr GJ, Chamberlain M, Curren RD. Validation of alternative methods for toxicity testing. *Toxicol in Vitro* 10:479-501 (1996).
15. Barratt MD. The role of structure activity relationships and expert systems in alternative strategies for the determination of skin sensitization, skin corrosivity and eye irritation. *ATLA* 23:111-122 (1995).
16. Whittle EG, Barratt MD, Carter JA, Basketter DA, Chamberlain M. The skin corrosivity potential of fatty acids: *in vitro* rat and human skin testing and QSAR studies. *Toxicol in Vitro* 10:95-100 (1995).
17. Oliver GJA, Pemberton MA, Rhodes C. An *in vitro* skin corrosivity test—modifications and validation. *Food Chem Toxicol* 24:507-512 (1986).
18. Whittle E, Basketter DA. The *in vitro* corrosivity test: comparison of *in vitro* human skin with *in vivo* data. *Toxicol In Vitro* 7:269-274 (1993).
19. EEC. Annexes I, II, III, IV to Commission Directive 93/21/EEC of 27th April 1993 adapting to technical progress for the 18th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Off J Eur Commun* L110A (1993).
20. Botham PA, Chamberlain M, Barratt MD, Curren RD, Esdaile DJ, Gardner JR, Gordon VC, Hildebrand B, Lewis RW, Liebsch M, et al. A prevalidation study on *in vitro* skin corrosivity testing. The report and recommendations of ECVAM Workshop 6. *ATLA* 23:219-255 (1995).
21. Archer GEB, Balls M, Barratt MD, Botham PA, Brantom PG, Curren RD, Earl L, Esdaile DJ, Holzhueter H-G, Liebsch M, et al. The ECVAM skin corrosivity validation study. *ATLA* 24:321 (1996).
22. Barratt MD, Castell JV, Chamberlain M, Combes RD, Dearden JC, Fentem JH, Gerner I, Giuliani A, Gray TJB, Livingstone DJ, et al. The integrated use of alternative approaches for predicting toxic hazard. The report and recommendations of ECVAM Workshop 8. *ATLA* 23:410-429 (1995).
23. ECETOC. ECETOC Technical Report No. 48, Eye Irritation: Reference Compounds Data Bank. Brussels: European Centre for Ecotoxicology and Toxicology of Chemicals, 1992.
24. Carpenter CP, Smyth HF. Chemical burns of the rabbit cornea. *Am J Ophthalmol* 49:1363-1372 (1946).
25. Potts RO, Guy R. Predicting skin permeability. *Pharmaceut Res* 9:663-669 (1992).
26. Barratt MD, Dixit MB, Jones PA. The use of *in vitro* cytotoxicity measurements in QSAR methods for the prediction of the skin corrosivity potential of acids. *Toxicol in Vitro* 10:283-290 (1996).
27. Basketter DA, Scholes EW, Chamberlain M, Barratt MD. An alternative strategy to the use of guinea pigs for the identification of skin sensitization hazard. *Food Chem Toxicol* 33:1051-1056 (1995).
28. Sanderson DM, Earnshaw CG. Computer prediction of possible toxic action from chemical structure: the DEREK system. *Human Exp Toxicol* 10:261-273 (1991).

29. Ridings JE, Barratt MD, Cary R, Earnshaw CG, Eggington E, Ellis MK, Judson PN, Langowski JJ, Marchant CA, Payne MP, et al. Computer prediction of possible toxic action from chemical structure; an update on the DEREK system. *Toxicology* 106:267–279 (1996).
30. Barratt MD, Basketter DA, Chamberlain M, Admans GD, Langowski JJ. An expert system rulebase for identifying contact allergens. *Toxicol in Vitro* 8:1053–1060 (1994).
31. Kimber I, Dearman RJ, Scholes EW, Basketter DA. The local lymph node assay: developments and applications. *Toxicology* 93:13–31 (1994).
32. EEC. Annex IV to Commission Directive 93/21/EEC of 27 April 1993 adapting to technical progress for the 18th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Off J Eur Commun* 36:59 (1993).
33. Basketter DA, Barratt MD, Chamberlain M, Griffiths HA, York M. Identification and classification of chemicals causing skin irritation: a strategy to replace animal tests. *ATLA* 24:233 (1996).
34. Barratt MD. A quantitative structure activity relationship (QSAR) for prediction of α -2 μ -globulin nephropathy. *Quant Struct Act Relat* 13:275–280 (1994).
35. Lake BG, Lewis DFV, Gra TJB, Beaman JA. Structure-activity relationships for induction of peroxisomal enzyme activities in primary rat hepatocyte culture. *Toxicol in Vitro* 7:605–614 (1993).
36. Ridings JE, Manallack DT, Saunders MR, Baldwin JA, Livingstone DJ. Multivariate quantitative structure-toxicity relationships in a series of dopamine mimetics. *Toxicology* 76:209–217 (1992).